INITIAL INVESTIGATION ON FATIGUE IN COMMAND AND CONTROL SITUATION AWARENESS: PHYSIOLOGY AND COGNITIVE PERFORMANCE

Johns Hopkins University
Applied Physics Laboratory

Jennifer McKneely
Matthew Bevan
Kevin Cropper
Mandy Iny
Frank Vaughan

11100 Johns Hopkins Rd
Laurel, MD 20723

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JUN 2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE **Initial Investigation on Fatigue in Command and Control Situation Awareness: Physiology and Cognitive Performance**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Johns Hopkins University,Applied Physics Laboratory,11100 Johns Hopkins Road,Laurel,MD,20723**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**15**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

ABSTRACT

Human decision making and situation awareness (SA) are critical in Command and Control (C2) making the C2 environment ideal for understanding implications of cognitive state on operational performance.

This pilot study investigated the effects of fatigue on SA, and the extent that physiological measures correlate with performance. Six novice participants were tested every three hours during the final 24 hours of a 36-hour period of wakefulness on SA, psychomotor vigilance (PVT), and Automated Neuropsychological Assessment Metrics (ANAM) tests. Additionally, heart rate, heart rate variability, and activity were monitored.

Initial analyses indicate changes in SA were not correlated with fatigue; potentially reflecting learning effects of fatigue, as opposed to SA. PVT results were similar to previous literature showing a significant performance drop in the early hours of the morning. The ANAM battery showed unexpected results; the Stanford Sleepiness scale correlated best (α=.01) with heart rate and second with activity (α=.05) and the logical reasoning-symbolic test result showed a statistically significant correlation with (α=.01) heart-rate variability.

A follow-on study will be conducted with experienced participants, (controlling the learning confound). Additionally, the study will investigate whether or not the degree of SA (high, medium, low) degrades differentially with fatigue.

INTRODUCTION

Human decision making and situation awareness (SA) are increasingly critical components to warfighting effectiveness, as is inferred in the attributes in the Cognitive Domain including “Awareness”’ “Understanding”, and “Sensemaking” in Figure 1 from the Department of Defense (DoD) Command and Control Research Program (CCRP), and others. Situation awareness can be hampered by large amounts of data and information that an individual must parse through in order to perceive what is going on, understand the data and how it is changing, notice trends and patterns, and predict what may happen in the future. To ensure maximum effectiveness and ability to maintain SA, warfighters must remain alert and cognizant of the environment around them.

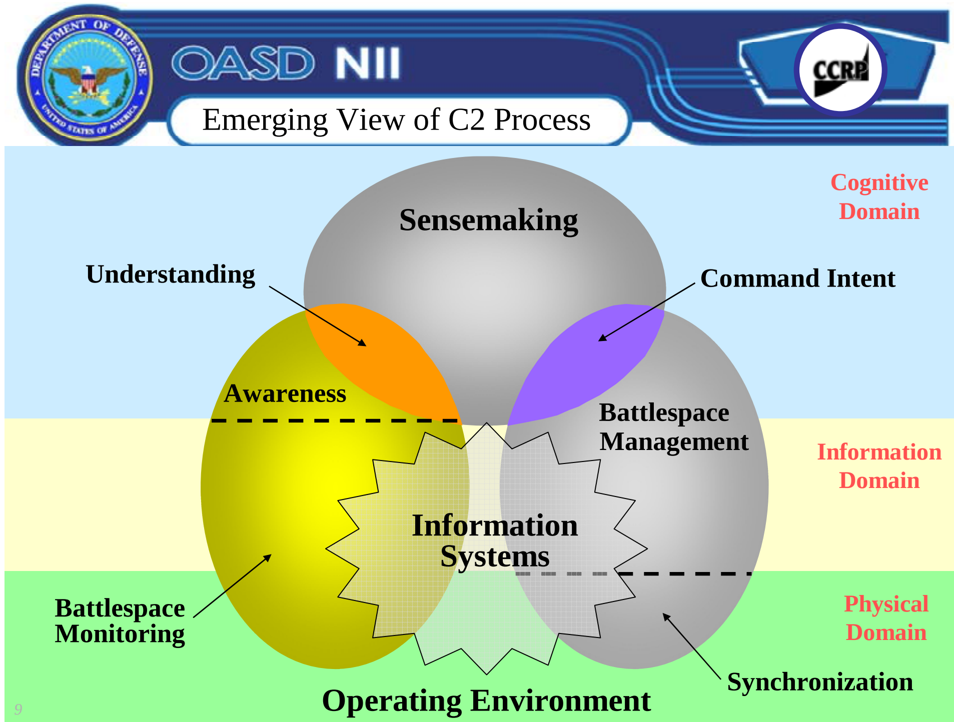

Figure 1 - Situation Awareness as a key aspect of Warfighter Effectiveness (from CCRP website)

Command and control (C2) is intense and demanding, requiring around-the-clock operation. Warfighters work extended shifts with small amounts of sleep between them. This can lead to a situation where decision makers are chronically fatigued. Chronic fatigue manifests in ways similar fatigue due to short term, yet significant sleep deprivation (36 hours or more of continued wakefulness). Research concerning pilots and others under conditions of sleep deprivation show decrements in performance in a wide array of tasks (vigilance, etc.) (Caldwell, Caldwell, Brown & Smith, 2004). This suggests that those same decrements in performance would be seen in C2 warfighters who are chronically fatigued. The decision making requirements of warfighters and the consequences of poor decisions demand that C2 personnel maintain alertness and cognitive functioning, at all times, even during those that are atypical.

Situation awareness often correlates with decision making, and is therefore a useful human performance metric to include in the evaluation of systems. However, numerous SA measures exist and have been used in military applications, they often include significant limitations. For example, some require a brief stop of the warfighter's activity to answer questions (e.g. Situation Awareness Global Assessment Technique, SAGAT) or are collected at the end of a task (e.g. Situational Awareness Rating Technique, SART) and provide only subjective measures. In addition, SA metrics, such as those mentioned and other similar SA measurements often do not result in practical measures that can be easily related and compared to other metrics. Therefore, we have derived an SA approach from Endsley's (2000) model of SA, complemented with a SA novel assessment method and metric tool that can be used in real-time, is comprehensive and flexible, increases the integrity of SA data collection and results in a single weighted

score. A useful function of our proposed SA metric is that it can be correlated with other metrics to evaluate relationships and interactions between various factors.

There is a need for Fatigue Countermeasures research and development. The DoD has documented mission failures and Class A mishaps, resulting in loss of trained personnel and valuable aerospace platforms due to human fatigue (Ramsey & McGlohn, 1997). Research has highlighted the potential negative effects on decision-making due to fatigue (Wilson, 2000). Fatigue effects have been shown to be similar to those of alcohol where after 28 hours of sleep an individual's performance correlated to that of an individual with 0.1% blood alcohol content (legally drunk in most states) (NRAC, 1999) (See Figure 2). Further, the effects of chronic sleep loss over extended times have been shown to decrease productivity substantially at the end of a three week period and effects can be seen as early as five days (NRAC, 1999) (Again, see Figure 2).

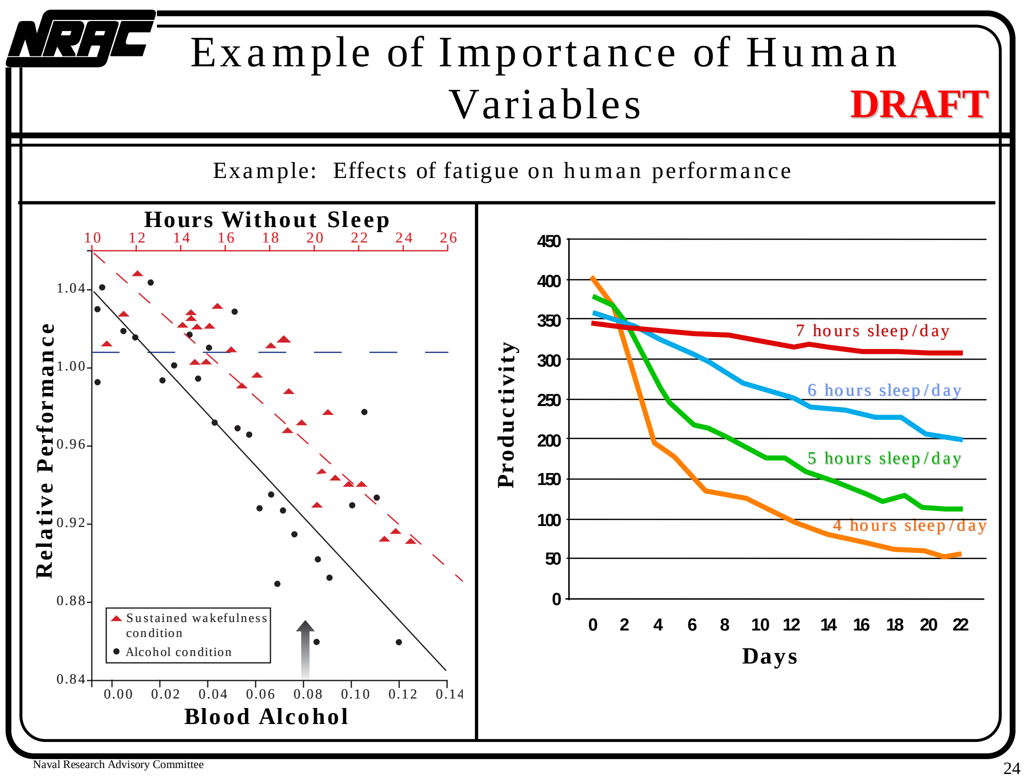

Figure 2 – Correlation between the effects of fatigue and alcohol and the effect of sleep on productivity.

However, the true nature of the potential performance decrement is not yet known. In addition, the mechanisms for assessing the level of fatigue need maturing. Physiological assessment and real time monitoring and feedback are possible; however, the predictive nature and utility of assessment during operations still needs validation. The relationship between physiological measures and cognitive performance must be firmly established prior to integration within C2 systems in operations.

In this pilot study, we demonstrated the ability to measure situation awareness over extended durations and tested a new SA principle, process, and measurement. Also, we began to investigate the relationship between fatigue (as measured with physiological sensors and the Automated Neuropsychological Assessment Metrics (ANAM) battery and the Psychomotor Vigilance Task (PVT)) and situation awareness.

We furthered our theory of SA for tactical environments and developed a framework to measure SA in real-time through research into the state-of-the-art in situation awareness, measurement of SA, and its relationship with fatigue. The tool we developed to measure SA was used in a pilot study evaluating six participants over a 36-hour duration.

Analysis of the data showed correlations between the established fatigue tests and the biometric sensors. Comparisons with the SA results were not as clear, but that may be explained by the novice abilities of the participants regarding the subject matter used for the experiment and an associated learning curve effect.

METHOD

Participants

Six adults (two males and four females) between the ages of 18 and 35 volunteered for the study in response to a recruitment e-mail. Based on self-reporting, participants did not have any chronic sleep disorders, neurological disorders, or other ailments that conflicted with their ability to perform tasks with extended periods of wakefulness. Each was instructed not to consume caffeine upon waking in the morning of the test though the end of testing. All of the participants gave written, informed consent and ensured safe transportation following the experiment was ensured. Approval for human participation was granted on 04 August 2005 by the Johns Hopkins Bloomberg School of Public Health (JHSPH) Committee on Human Research Institutional Review Board.

Test Components & Materials

- 2 Personal Computers running the following software:
  - Automated Neuropsychological Assessment Metrics (ANAM) Software developed by the U.S. Army Medical Research and Materiel Command
  - Psychomotor Vigilance Task (PVT) Software for recording data from PVT
  - SA Task- simplified Tactical Tomahawk Weapon Combat System (TTWCS) v.6 prototype run on PowerPoint software
  - Matlab Data Acquisition Software
- Biometric Sensors
  - Electrocardiogram (ECG) to measure heart-rate and heart-rate variability
  - Microwave Sensor System to measure physical activity, slow eye closure, blink length, heart-rate variability,
- Informed consent package (signature forms and experimental overview)
- Experimenter Script
- Psychomotor Vigilance Task (PVT) Hardware
- SA Script and Probes
- SA Record Sheets for test administrators

- Modified Situational Awareness Rating Technique (SART)
- Experimenter Log
- Entertainment for participants during off-test periods– movies, hand held games, music, cards, etc.

Procedure

Two test sessions were held with three participants tested individually for eight test cycles during each session. A training session was conducted the evening prior to the test session.

During training, an informed consent script was read to all of the participants, which included an overview of the study. Familiarity training was provided on the PVT, ANAM, and TTWCS SA tasks. Participants had to achieve 85% accuracy on four ANAM tests (Code Substitution Learning, Math Processing, Logical Reasoning Symbolic, Run CPT). The participants were also shown how to place the ECG electrode patches and connect the leads. The ANAM Test Battery- consisted of seven tests:

- Sleep scale-R (Stanford Sleepiness Scale)
- Code Substitution Learning
- Math Processing
- Logical Reasoning Symbolic
- Run CPT (Complex Attention Measure)
- Code Substitution Delayed
- Sleep scale-R (Stanford Sleepiness Scale)

Participants received a wake-up call two hours before their baseline start time and were instructed to arrive at the test site at staggered test times (8 a.m., 9 a.m., or 10 a.m.) to establish baseline readings and review the training. Baselines and all trials were conducted on each participant with all tasks being performed while collecting physiological data; each test cycle took approximately thirty minutes. The tasks began with a two minute quiet period to stabilize the physiological data, followed by performing PVT for five minutes, ANAM for 15-20 minutes, and TTWCS SA task and modified SART for approximately 10 minutes. At the conclusion of the baseline run, participants continued with their normal work/leisure activity without consuming alcohol and caffeine or taking a nap.

After spending the day conducting their usual daily activities, participants returned to the testing location ten hours after their baseline test, at 6 p.m., 7 p.m., or 8 p.m., respectively. Every three hours over the next twenty-four hours, testing was conducting repeatedly. Participants were monitored for the experiment duration, were monitored for wakefulness, and provided with entertainment and food/beverages (no caffeine) during testing. The investigation team comprised three pairs of investigators, each working one nine-hour shift of the twenty-four hour testing period allowing for overlap to ease with transition.

Data Collection

The independent measure was number of hours of wakefulness. Dependent measures included heart rate, heart rate variability, and physical activity, psychomotor vigilance task (PVT) reaction time, Automated Neuropsychological Assessment Metrics (ANAM) test scores, and SA scores.

JHU/APL has developed an enhanced situation awareness definition and measurement tool (Provisional Patent 2198-6606). The definition is based on Endsley's (2000) definition of SA (level 1=perception, level 2=comprehension, level 3=projection) to which JHU/APL has added a level between Endsley's level 2 and level 3 for trend analysis. Hence, the JHU/APL SA definition utilized was separated into 4 levels:

1. Perception,
2. Comprehension,
3. Analysis,
4. Projection.

To measure SA, SA assessment/probe questions applicable to the TTWCS task were developed and assigned a level of SA (selected examples in Table 2). While participants performed the SA task, the test administrator recorded the participants' responses and judged the participants' confidence. Participants also subjectively assessed their confidence with the modified version of the Situational Awareness Rating Technique (SART). The SART typically includes ratings of the supply and demand of attention and understanding; the modified version also included a rating of self-confidence.

Probe Question	**SA Level**
What is the estimated first time of launch?	*1*
Are there any tracks that might interfere with the OTW route?	*1*
What is the time of impact for CFF 1?	*2*
Can you still fire this missile, why/why not?	*2*
How long until the last BLK III D has its data loaded?	*3*
What time will you have your next health and status message event?	*3*

Table 2 – Example Situation Awareness probes, with assigned SA Level

RESULTS AND DISCUSSION

The tests consisted of collecting biometric sensor (ECG and microwave) data while participants completed PVT, ANAM, and SA/SART tasks. Comparisons were conducted on a sub-set of the tasks that seemed most appropriate to command and control. The data from all participants were averaged within each cycle. Correlations between these averaged measures were assessed, and Table 3 details the Pearson Correlation Coefficients, the probability it would result under a null hypothesis with r=0.95, and the number of observations used in the comparison (some measures did not have data in a particular cycle and so had fewer observations than others), Figure graphs the averages. Some correlations were strong enough to hint at a potential link between the two measures and support further investigation in future studies. More detailed descriptions of the data collection and analysis methods for each of the tests (biometric sensors, PVT, ANAM, and SA) follow this general comparison.

Table 3: Correlation Coefficients, Probability, and Number of Observations for a subset of the collected datasets. Coefficients < 0.5 are highlighted in pale yellow; coefficients < 0.9 are highlighted in tan. Probabilities stronger than 0.1 are highlighted in light turquoise; probabilities stronger than 0.05 are highlighted in blue; probabilities stronger than 0.01 are highlighted in purple; and probabilities stronger than 0.001 are highlighted in maroon.

Pearson Correlation Coefficients; Prob > |r| under H0: Rho=0; r=0.95
Number of Observations (0 to 9; baseline, cycle 1, cycle 2, …, cycle 8)

	SA 1	SA 2	SA 3	PVT Average Delay	PVT Sleepy	Heart Rate	Heart Rate Var.	ANAM Sleepy	ANAM Math	ANAM Logical	ANAM Code
Overall SA	0.814 0.026 7	0.921 0.003 7	0.359 0.484 6	-0.165 0.696 8	-0.514 0.192 8	0.193 0.647 8	0.006 0.990 8	-0.560 0.149 8	0.325 0.432 8	0.124 0.770 8	0.338 0.414 8
SA 1		0.579 0.229 6	0.369 0.541 5	0.181 0.698 7	-0.275 0.550 7	0.020 0.967 7	0.204 0.661 7	-0.320 0.484 7	-0.037 0.937 7	-0.280 0.543 7	0.021 0.964 7
SA 2			0.278 0.650 5	-0.255 0.581 7	-0.492 0.263 7	0.199 0.668 7	-0.102 0.828 7	-0.620 0.138 7	0.567 0.185 7	0.228 0.624 7	0.499 0.254 7
SA 3				-0.254 0.627 6	-0.446 0.376 6	0.512 0.300 6	-0.143 0.787 6	-0.201 0.702 6	0.333 0.519 6	0.610 0.199 6	-0.279 0.592 6
PVT Average Delay					0.773 0.015 9	-0.700 0.036 9	0.693 0.038 9	0.750 0.020 9	0.168 0.665 9	-0.818 0.007 9	-0.614 0.079 9
PVT Sleepy						-0.888 0.001 9	0.720 0.029 9	0.953 <.0001 9	0.580 0.102 9	-0.729 0.026 9	-0.661 0.053 9
Heart Rate							-0.744 0.022 9	-0.785 0.012 9	-0.554 0.122 9	0.640 0.063 9	0.369 0.329 9
Heart Rate Var.								0.752 0.019 9	0.311 0.415 9	-0.717 0.030 9	-0.535 0.138 9
ANAM Sleepy									0.487 0.184 9	-0.641 0.063 9	-0.817 0.007 9
ANAM Math										-0.147 0.705 9	-0.319 0.404 9
ANAM Logical											0.460 0.213 9

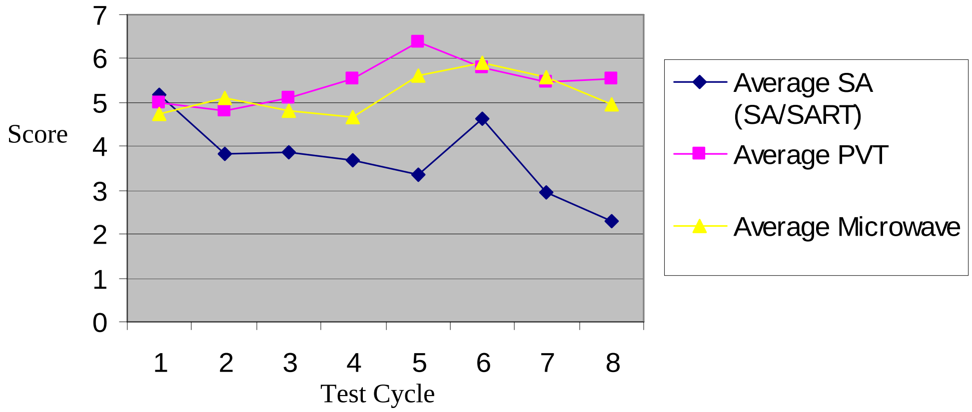

Figure 3, Average SA, PVT, and Microwave by test cycle

The Situation Awareness measures showed some strong internal correlations, as the Overall SA, SA Level 1, and SA Level 2 all were above 0.5. It is possible that Overall SA could be sufficient as a single summary of SA, so measuring individual levels would be unnecessary. However, the SA Level 3 did not correlate with the other SA measures which was not surprising as participants were novice users and did not possess the knowledge of the subject area or system to make accurate projections about future trends.

The PVT measure represented fatigue in this study; therefore, the strong correlations it exhibited with the measures other than SA provides encouragement that ANAM or Heart Rate could also be used to successfully measure fatigue. The self-reported ANAM Sleepy and PVT Sleepy scores correlated significantly. The PVT Average Delay scores also correlated strongly which was supported by anecdotal comments from the Experimenters who observed participants reactions to be much slower when tired.

PVT Sleepy scores correlated positively with Heart Rate Variability, ANAM Sleepy, and ANAM Math. While the first two seem reasonable, the latter is an oddity that warrants further investigation. Why would math scores increase as self-reported sleepiness increased? Several possible explanations have been discussed including that math is an automated process for the participants in this study (engineering students) and that math test was not interesting under normal states, but due to its ease, was found to be more interesting in fatigue states.

Both PVT measures had negative correlations with Heart Rate, ANAM Logic, and ANAM Code, which would be expected. The more tired the participant, the longer the delay for the PVT, the slower the heart rate, and the less capable of solving logical problems.

Biometric Sensors

Two different biometric sensors were used in this study. The microwave sensor, positioned over the computer monitor, detected movements of the participants, from broad head nodding, to eye closure, and possibly down to the level of pulsing blood

vessels, yielding a measure of heart rate. The ECG, attached to the participants directly, recorded heart rate in a more traditional manner.

Microwave Activity

The microwave activity data were reduced to provide a numerical measure of relative activity for each test sequence of each cycle for each of the test subjects. The results were then analyzed graphically to determine if there was a correlation between the activity metric and time without sleep. In earlier work, activity had been shown to decrease with the onset of fatigue. However, in this "keep awake" test an opposite trend was seen.

The microwave data file was reduced by calculating the average for the period of the test. For the sitting and PVT portions of the test, the test length was precisely known. For the ANAM and Situation Awareness tests, the test time varied with the speed of the subject's responses. Since the data collection period was specified at the beginning of the test and needed to be longer than the anticipated test period, the test file was truncated based on entries in the test logbook that listed the start and stop times (to the closest minute) for these tests. To avoid incorporating excessive time, the last minute of data was truncated. This method of truncating data files was chosen because budget limitations did not allow each of the approximately 108 ANAM and Situation Awareness data files to be individually plotted, examined and truncated.

The average was calculated by squaring and summing the vertical and horizontal polarizations of the Doppler microwave signal measurements collected at 1 kHz. The square root of this value was calculated. This square root value was averaged over the data collection period. This resulted in an average measurement that reflected the aggregated movement of the person in front of the sensor. Figure 4 graphs the average measurement by participant across cycles.

There are several possible explanations for this finding, including:

- The microwave sensor used in the testing had a very narrow beam pattern requiring careful alignment of the device to aim at the head of the subject. It is suspected that the abnormally low measurements made in the first series of tests were due to the microwave device not being properly aimed at the subject.
- The tasks required concentration on a small screen in a monotonous environment. This naturally encouraged a focused effort suppressing idle fidgeting. This would naturally reduce activity in contrast with tasks where low levels of concentration were required or there is natural motion associated with the task such as in the drowsy driver application in which this sensor was previously employed.
- The subjects were kept awake and alert to the testing. While someone is alert and responsive, activity is maintained. In fact the increase in activity might reflect a natural countermeasure to fatigue, i.e. to increase activity to stay awake.

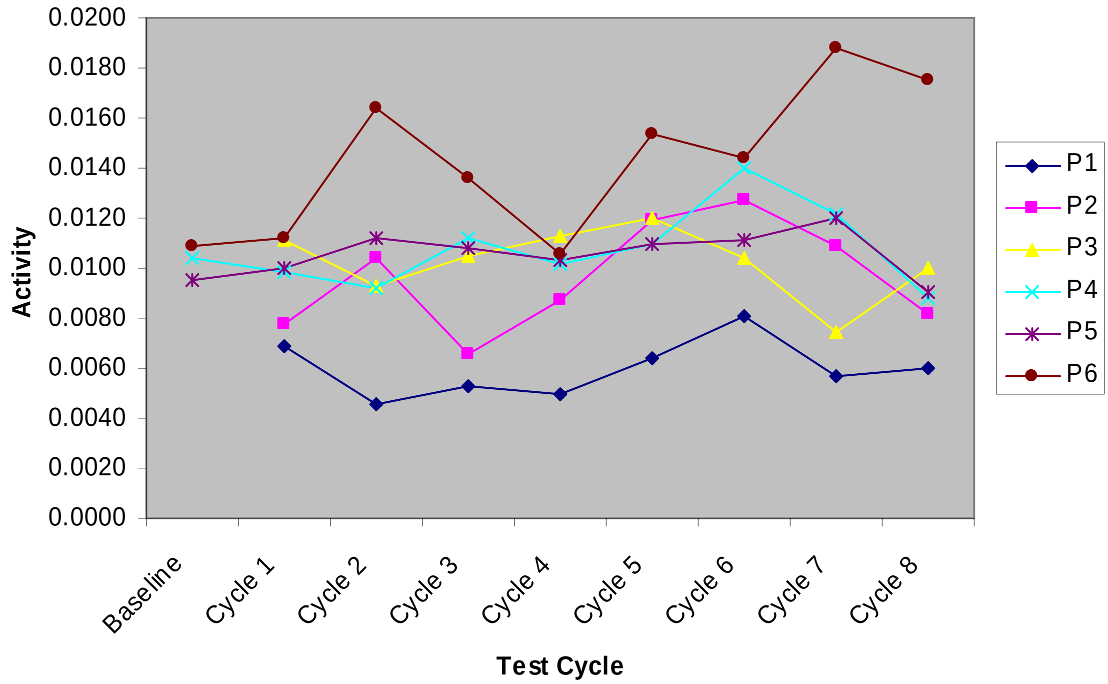

Figure 4 – Data showing the activity for each participant during testing. Activity reflects the average of the Doppler microwave signal during the four test periods for each cycle.

Electrocardiogram (ECG)

The Heart Rate and Heart Rate Variability scores are difficult to draw conclusions from and directly tie to ANAM scores, as the data analyzed were recorded at the beginning of the testing cycle, not simultaneously with the other tests. However, some interesting results were indicated, and certainly future studies should investigate the Heart Rate measures in closer temporal proximity to the other measures if possible. Changes to the apparatus would likely be necessary, as the Heart Rate measures were affected by the movement of the participants, even motions as small as typing at a keyboard or pressing a single key. These muscle movements created a significant amount of "noise" in the measurements, and a way to avoid or clean them would be necessary.

PVT

The data collected for the participants included their self-reported Sleepy Rating (on a 1 to 8 scale) before and after running the PVT task, as well as the delay (in milliseconds) for responding to each event and the event start time (in tenths of seconds) after the task began. Within the data, some anomalous results were seen. The data sets for several participants included impossible (0.001 seconds) or questionable (under 0.100 seconds) reaction times. Testing with the PVT equipment showed that times of 0.001 were recorded when the participant made a false report (pressed the button before the counter started); this was probably due the finger slipping as the participant was nodding off or jerking awake. It was hypothesized that the questionable times (the three abnormal responses were actually 3ms, 47ms, and 51ms) were recorded when the participant made

a false report just as the timer began counting. For this data analysis, both ‘impossible’ and ‘questionable’ data points were removed. Other data culled from the data sets included maximum delay (nearly three seconds) and number of delays over 400 milliseconds (several around 20).

During Test Session 1, there was a battery malfunction in the PVT equipment, rendering it impossible to collect data from every participant during every cycle. This analysis used averages based upon all available data recordings, not just those of participants 4-6 (Test Session 2). Thus, in some cycles, the PVT scores used in the correlation comparison are the average of just 3 (and occasionally 4) participants.

The data indicate a notable increase in response times during cycle 5, at which point the participants would have been awake for 24 hours straight (the 6, 7, and 8 a.m. testing periods). This corresponds to the highest sleepiness indication reported by the participants.

ANAM

Within the ANAM battery, the self-reported Sleepy score provided an indication of a participant’s performance on the Logic and Code sections, as scores decreased as the sleepiness increased. Sleepiness indications were given before and after the PVT and ANAM tests which may influence response…when the participant said, “I’m sleepy”, did that prime the brain to respond more slowly? or did the participant think, “I felt pretty slow there, so I must be sleepy”?

SA/SART

In addition to suggesting that future studies might benefit from experienced or expert users in order to combat learning and inexperience effects, the strength of the SA results indicate that the breadth and depth of the SA probes could be improved as well. There were no SA Level 4 probes, several of the cycles had no probes at one of the other SA Levels, and many had only one or two probes at a Level. Thus a single, possibly lucky, high or low score could imbalance a cycle and throw off correlations with other measures. Ensuring that several probes of each level are asked during each cycle would provide more robust results.

In comparisons to the other measures, SA had only moderately consistent results. The ANAM Sleepy and PVT Sleepy correlations were all negative, as would be expected; increases in the participant’s self-reported sleepiness led to decreases in SA scores. This was most consistently pronounced for the Overall SA.

The number of SA Level 2 probes that required participants to do math (adding remaining time to current time) might explain the correlation with ANAM Math. Similarly for the SA Level 3 correlation with the ANAM Logic, as logical thinking was necessary to make the predictions for Level 3 probes. However, the SA Level 3 correlation with Heart Rate (HR) cannot be attributed to an increase in stress due to the difficult questions, as the measures were not made simultaneously.

DISCUSSION

This pilot study has demonstrated that the testing protocol supports assessment of Situation Awareness (SA) of fatigued individuals. As is common in the fatigue literature, we found a decrement in concentration (as measured by PVT) around 24 hours of wakefulness; however, we did not find a significant decline in SA. This is postulated to be the result of testing novice (summer interns) users as opposed to experienced Tomahawk warfighters.

We have also demonstrated effects of fatigue on activity (Figure 3) where the correlation of fatigue with cycles suggests that the longer people are awake, the greater their activity. Although weak, there appears a trend with 5 of the 6 subjects of increasing activity with increasing time without sleep. In comparing these results with the PVT results, the PVT results showed a marked change in concentration in Cycles 5 and 6 that was not observed in the microwave activity measurement. The activity monitor is intended to detect the onset of sleep by measuring the change in activity levels; in this study it did not show sensitivity in detecting the lack of sleep or loss of attention as measured by the PVT test. In tasks that require focused concentration with the subject constrained to be sitting in front of and viewing a fixed display, the range of activity between awake and asleep states is reduced thus further reducing its sensitivity. While the microwave activity sensor may be well suited to detect the onset of sleep, this testing does not show it to be effective in detecting the lack of sleep in this type of physical configuration.

The ANAM battery effectively measured cognitive processes that are similar to those employed in command and control tasks, the logic and code substitution tasks correlated with PVT scores indicating that the more tired the individual the less capable of solving cognitively demanding tasks.

The DARPA Augmented Cognitive program is exploring a variety of sensors to gauge cognitive states of fatigue, workload, and stress. There are promising research findings regarding the ability of eye trackers, electro-encephalograms, and electrocardiograms to objectively measure fatigue. The employment of additional sensors may provide better measures of fatigue either through individual sensitivity to fatigue or to a composite measure based on multiple sensors.

A follow-on experiment is planned for Summer 2006. The experiment will use the protocol presented in this paper. To better gauge the effect of fatigue on SA we will evaluate Subject Matter Experts by testing current Fleet operators. This warfare domain will be undersea warfare supporting the generalization of fatigue effects across warfare areas. In addition to testing Fleet operators, we will also employ eye tracker to analytically evaluate scanning patterns and measure Fatigue.

It is recognized SA is influenced not only by cognitive state; but, also by system design. To enrich our evaluation of SA, we will compare two instantiations of a sonar system to determine if different qualities of SA are achieved. One will have a visual and single-

channel aural display and the second will have a visual and 3D spatial aural display. It is hypothesized that the second system configuration will provide a better initial SA and this benefit will be maintained over the 36-hour period of wakefulness.

It is expected that the follow-on study, combined with the results of this current pilot study, will provide beneficial data on how situational awareness is affected by fatigue giving system designers and policy planners insight into human-machine interface design and team scheduling.

ACKNOWLEDGMENTS

Several individuals provided assistance with staffing of the observation teams. Dr. Jennifer Ockerman and Dr. Kendal Roberts both gave generously of their weekend to support the study. The team would also like to thank the Volpe National Transportation Systems Center for the loan of the PVT device for our experiments.

REFERENCES

http://www.dodccrp.org/

Bevan, M., M. Dykton, J. A. B. McKneely, and P. Schuster, (2005). Linking Fatigue Sensing to Operational Performance: Using Human-Systems Integration Methods to Demonstrate the Operational Value of Fatigue Sensors, JHU/APL Internal White Paper.

Caldwell, J. A., J. L. Caldwell, D. L. Brown, and J. K. Smith, (2004) "The Effects of 37 Hours of Continuous Wakefulness on the Physiological Arousal, Cognitive Performance, Self-Reported Mood, and Simulator Flight Performance of F-117A Pilots," Military Psychology, 16(3), 163̃181.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Naval Research Advisory Committee (NRAC) (1999), Optimized Surface Ship Manning.

Marshall, Sandra P. (2005). Assessing Cognitive Engagement and Cognitive State from Eye Metrics, Paper presented at the 1st Annual Conference on Augmented Cognition, Las Vegas, NV, July 22-27, 2005.

Ramsey, X., and Y. McGlohn, (1997). Zolpidem as a fatigue countermeasure. Aviation, Spcace and Environmental Medicine. 68(10), 926-931.

Wilson, G. F. (2000) "Strategies for Psychophysiological Assessment of Situation Awareness," in Situation Awareness Analysis and Measurement, Lawrence Erlbaum Associates, Mahwah, NJ.